

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Wolf, Marcia
Cassels, Frederick
Boedeker, Edgar
- (ii) TITLE OF INVENTION: Transformed Bacteria Producing CS6
Antigens and Vaccines

(iii) NUMBER OF SEQUENCES: 8

- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Glenna Hendricks
(B) STREET: P.O. Box 2509
(C) CITY: Fairfax
(D) STATE: VA
(E) COUNTRY: USA
(F) ZIP: 22031

- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US 09/479,877
(B) FILING DATE: 10-JAN-2000
(C) CLASSIFICATION:

- (vii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Hendricks, Glenna
(B) REGISTRATION NUMBER: 32,535
(C) REFERENCE/DOCKET NUMBER: army 09/479,877

- (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (703) 425-8405
(B) TELEFAX: (703) 425-8406

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4876 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: E. coli
- (B) STRAIN: CS6

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AAGCTTGTAA CCAGTTGATA AAAATATATC ACGCTGGAA TGACGTGATG TATATACGGA
60

GCAGCTATGT CGGAACAGAT ATTTTCCTAT CGGTATGCGT TGTGAGTAAG CGTAAAGCCA
120

ATGCTGTCTG TAACTCCTGA TCCTTGCAGA CTAAATTAGA GCTCCTTCTA AATTAGACGG
180

ATGGATAAAC CTACAGACTG GCGCTCTGGG TCTGCCCGGA TATTTTCTAA TGAATTAAAG
240

CTTCATATGG TTGAACCTGGC TTGAAACCCA AATGCCAATG TCGCACAACT GGCTCGGGAA
300

CATGGCGTTG ATAACAACCT GATTTTAAA TAGCTACGCC TCTGGCAAAG AGAAGGACGT
360

ATTTCTCGTA GAATGCCTCC AACTATTGTA GGCCCTACAG TACCACTGAG GTAGCCTGAA
420

TTTAAAGCCG AAGCGGTCAG AACTGTTCTT GGTGTGAAACG TAGCACTCAC CAATAAAAGC
480

ATCAATACGG TGCTCTGTG ACACATTACG AATGTTATGT ATACAATAAA AATGATTATA
540

GCAATATTAA TGGTGTATA TGAAGAAAAC AATTGGTTA ATTCTAATTG TTGCTTCATT
600

CGGCAGCCAT GCCAGAACAG AAATAGCGAC TAAAAACTTC CCAGTATCAA CGACTATTC
660

AAAAAGTTT TTTGCACCTG AACCACGAAT ACAGCCTCT TTTGGTGAAA ATGTTGGAAA
720

GGAAGGAGCT TTATTATTAA GTGTGAACTT AACTGTTCTT GAAAATGTAT CCCAGGTAAC
780

GGTCTACCCT GTTTATGATG AAGATTATGG GTTAGGACGA CTAGTAAATA CCGCTGATGC
840

TTCCCAATCA ATAATCTACC AGATTGTTGA TGAGAAAGGG AAAAAAATGT TAAAAGATCA
900

TGGTGCAGAG GTTACACCTA ATCAACAAAT AACTTTAAA GCGCTGAATT ATACTAGCGG
960

GGAAAAAAA ATATCTCCTG GAATATATAA CGATCAGGTT ATGGTTGGTT ACTATGTAAA
1020

CTAAATACTG GAAGTATGAT TATGTTGAAA AAAATTATTT CGGCTATTGC ATTAATTGCA
1080

GGAACCTCCG GAGTGGTAAA TGCAGGAAAC TGGCAATATA AATCTCTGGA TGTAAATGTA
1140

AATATTGAGC AAAATTATTC TCCAGATATT GATTCCGCTG TTCTGTATAAT ACCTGTTAAT
1200

TACGATTCCG ACCCGAAACT GGATTCACAG TTATATACGG TTGAGATGAC GATCCCTGCA
1260

GGTGTAAAGCG CAGTTAAAAT CGCACCAACA GATAGTCTGA CATCTCTGG ACAGCAGATC
1320

GGAAAGCTGG TTAATGTAAA CAATCCAGAT CAAAATATGA ATTATTATAT CAGAAAGGAT
1380

TCTGGCGCTG GTAACCTTAT GGCAGGACAA AAAGGATCCT TTCCTGTCAA AGAGAATACG
1440

TCATACACAT TCTCAGCAAT TTACTGGT GGCGAATACC CTAATAGCGG ATATTCTCT
1500

GGTACTTATG CAGGAAATTT GACTGTATCA TTTTACAGCA ATTAAAAAAA GGCCGCATTA
1560

TTGCGGCCAT TGACGATACT GCTAGGCAAA AATATGAAAT CAAAGTTAAT TATACTATTG
1620

ACGTTAGTGC CATTTCATC TTTTCAACA GGAAATAATT TTGAAATAAA TAAGACACGA
1680

GTAATTTACT CTGACAGCAC ACCATCAGTT CAAATATCAA ATAATAAAGC ATATCCTTTA
1740

ATTATTCAAA GCAATGTATG GGATGAAAGC AATAATAAAA ATCATGACTT TATAGCAACA
1800

CCACCGATTT TTAAAATGGA AAGTGAAAGT CGGAATATAA TAAAAATAAT TAAAACAAC
1860

ATTAATTGCG CGGACTCTCA GGAAAGTATG AGATGGTTAT GTATTGAATC AATGCCACCA
1920

ATAGAAAAAA GTACTAAAAT AACAGAAAA GAAGGAAGGA CAGACAGTAT TAATATCAGC
1980

ATTCGGGGGT GCATTAAACT GATATATCGA CCTGCCAGTG TTCCGTCTCC TGTTTTAAT
2040

AATATAGTAG AAAAATTAAA ATGGCATAAA AATGGAAAGT ATCTTGTATT AAAAAATAAT
2100

ACACCCTATT ACATTAGCTT TTCTGAGGTT TTTTTGATT CAGATAAAGT AAACAATGCA
2160

AAAGATATTT TATATGTAAA ACCATACTCA GAGAAGAAAA TAGATATCAG CAACAGAATA
2220

ATAAAAAAA TCAAATGGGC TATGATTGAT GATGCTGGCG CAAAAACAAA ACTTTATGAA
2280

TCAATTATAT AAAAATCTC ATTACAGTAT ACAAAAACAT CAGATTACAG GCTTGCTTTT
2340

TTTGCCTTTT ATATATCCTT TCTCACCTC ATATGGAAAT GAACAATTAA GTTTGACTC
2400

ACGATTCCCTA CCATCAGGTT ATAATTACTC TTTAAATAGT AACTTACCTC CTGAAGGTGA
2460

GTATCTGGTT GATATTATA TTAACAAAAT AAAAAGGAG TCCGCGATTA TTCCTTTTA
2520

TATAAAAGGA AATAAACTTG TACCATGTTT ATCAAAAGAA AAAATTCAT CTTGGGTAT
2580

CAACATTAAT AATAACGACA ACACAGAGTG TGTAGAAACA AGTAAGGCAG GTATTAGTAA
2640

TATCAGCTTT GAGTTTAGCT CTCTCGTTT GTTTATTGCT GTACCGAAAA ATCTTCTGTC
2700

TGAGATTGAT AAAATATCAT CAAAGGATAT AGATAACGGG ATTGATGCTT TATTTTTAA
2760

TTATCAAGTA AATACAAGGC TAGCCAATAA TAAAAATCGT TATGATTACA TTTCTGTTTC
2820

ACCAAATATA AATTATTTT CATGGCGGTT GCGTAATCTT TTTGAATTAA ACCAAAACAA
2880

CGATGAAAAA ACATGGAAA GAAACTACAC TTATCTAGAA AAAAGTTTT ATGATAAAAA
2940

GCTAAACTTA GTCGTTGGTG AAAGTTATAC GAATTCAAAT GTTTATAATA ACTACTCTT
3000

TACTGGTATT TCAGTTCTA CAGATACAGA TATGTATACG CCAAGTGAAA TCGATTATAC
3060

ACCAAGAAATT CATGGAGTGG CTGATTCAGA CTCTCAGATT ATTGTCAGGC AAGGCAACAC
3120

CATTATCATT AATGAAAGTG TTCCAGCCGG ACCGTTCTCA TTTCCAATAA CCAATCTCAT
3180

GTATACTGGG GGGCAACTTA ATGTGGAGAT AACAGATATT TATGGAAATA AAAAACAAATA
3240

TACTGTCAAT AATTCCCTCTC TTCCTGTTAT GAGAAAAGCG GGACTAATGG TATATAATTT
3300

TATATCTGGG AAATTAACAA AAAAAAATAG TGAGGATGGT GATTTTTTA CTCAAGGTGA
3360

TATTAACTAC GGTACTCACT ATAACAGCAC ACTATTGGT GGATATCAGT TTAGTAAAAA
3420

TTATTTAAC TTATCTACTG GTATAGGCAC TGATCTGGGA TTTTCTGGAG CATGGCTACT
3480

ACACGTTAGC AGAAGTAATT TTAAGAATAA AAATGGATAT AATATTAATC TACAACAAAA
3540

CACTCAGTTA AGACCATTCA ATGCCGGGGT TAATTCGAT TACGCATACA GAAAAAAAAG
3600

GTATGTGGAA CTTTCCGACA TTGGCTGGCA TGGTAATTAA TATAATCAAC TTAAAAATAG
3660

TTTTCTTTA TCCTTGTCAA AATCATTGAA TAAATACGGA AATTTCTCAC TTGATTATAA
3720

CAAAATGAAA TACTGGATA ATGCGTATGA TAGTAACTCA ATGTCGATTC GTTATTTTT
3780

TAAATTCAATG CGAGCAATGA TTACAACAAA TTGTTCTTTA AATAAATATC AATCTTATGA
3840

AAAAAAAGAT AAAAGATTAA GTATTAATAT ATCATTGCCT TTAACCAAAG ATTACGGGCA
3900

CATATCTTCA AACTATTCAAT TTTCCAATGC AAATACAGGA ACAGCAACCA GTTCTGTAGG
3960

CTTAAACGGT AGTTTTTTA ATGACGCAAG ATTAAACTGG AACATTCAAGC AGAACAGAAC
4020

GACCCGTAAC AATGGATATA CTGATAATAC CAGTTACATA GCAACCAGCT ATGCCTCTCC
4080

CTATGGCGTT TTTACTGGTT CATATTCAAGG ATCGAACAAAG TATTCAAGCC AGTTTTATTC
4140

TGCATCGGGA GGTATTGTTT TGCATAGCGA TGGCGTAGCT TTTACTCAA AAGCCGGAGA
4200

TACCTCTGCT CTTGTCCGTA TTGATAATAT TTCTGATATA AAAATTGGTA ACACTCCTGG
4260

TGTTTATACT GGGTATAATG GTTTGCTTT AATTCCCAT CTTCAGCCGT TCAAAAAAAA
4320

CACCATTTA ATTAATGATA AAGGAATTCC AGACGGTATT ACTCTTGCTA ATATAAAAAAA
4380

ACAAGTTATC CCATCACGAG GAGCTATTGT TAAAGTAAAA TTTGATGCTA AAAAAGGCAA
4440

TGACATTTG TTTAAGCTTA CAACTAAAGA TGGAAAAACG CCCCCATTAG GAGCTATAGC
4500

CCATGAAAAA AATGGAAAAC AGATTAATAC GGGTATCGTT GACGATGATG GTATGCTTTA
4560

TATGTCTGGA TTATCAGGGA CAGGGATTAT TAATGTAACA TGGAATGGAA AAGTCTGTTC
4620

ATTCCTTT TCAGAAAAAG ATATATCTAG CAAACAATTA TCTGTTGTAATAAACAATG
4680

TTAGGTAGTG CATCCAATTA GTAGAACATG TGTTTTCGA TAAACGCTCC GATCTCTTT
4740

TCGTGGATCT CAACTGAGCG TGAGAACCGAG ATTGTTTAC GAGCCAACCG CTTAATGCGG
4800

GTCGCTAGCG TCAGATTATT ACGCTCAATG CGTTGGGTGA ATATTTGCC GGTCAGATGC
4860

TTATTCTTCG GTACCC
4876

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 18 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: E. coli
(B) STRAIN: CS6

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GGCCGCATTA TTGCGGCC
18

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

- (vi) ORIGINAL SOURCE:
(A) ORGANISM: E. coli
(B) STRAIN: CS6

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GGCCGCATTA TTGATTGCGG CC
22

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 527 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

- (vi) ORIGINAL SOURCE:
(A) ORGANISM: E. coli
(B) STRAIN: CS6

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

TTGACACATT ACGAATGTTA TGTATAACAAT AAAAATGATT ATAGCAATAT TAATGGTGT
60

ATATGAAGAA ACAATTGGT TTAATTCTAA TTCTTGCTTC ATTCCGCAGC CATGCCAGAA
120

CAGAAATAGC GACTAAAAAC TTCCCAGTAT CAACGACTAT TTCAAAAAGT TTTTTGCAC
180

CTGAACCACG AATACAGCCT TCTTTGGTG AAAATGTTGG AAAGGAAGGA GCTTTATTAT
240

TTAGTGTGAA CTTAACTGTT CCTGAAAATG TATCCCAGGT AACGGTCTAC CCTGTTATG
300

ATGAAGATTA TGGTTAGGA CGACTAGTAA ATACCGCTGA TGCTTCCCAA TCAATAATCT
360

ACCAGATTGT TGATGAGAAA GGGAAAAAAA TGTTAAAAGA TCATGGTGCA GAGGTTACAC
420

CTAATCAACA AATAACTTT AAAGCGCTGA ATTATACTAG CGGGGAAAAA AAAATATCTC
480

CTGGAATATA TAACGATCAG GTTATGGTTG GTTACTATGT AAACTAA
527

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 154 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: N-terminal

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: E. coli
- (B) STRAIN: CS6

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Ser Met Lys Lys Thr Ile Gly Leu Ile Leu Ile Leu Ala Ser Phe Gly
1 5 10 15

His	Ala	Arg	Thr	Glu	Ile	Ala	Thr	Lys	Asn	Phe	Pro	Val	Ser	Thr
Thr														
	20						25						30	
Ile	Ser	Lys	Ser	Phe	Phe	Ala	Pro	Glu	Pro	Arg	Ile	Gln	Pro	Ser
Phe														
	35						40					45		
Gly	Glu	Asn	Val	Gly	Lys	Glu	Gly	Ala	Leu	Leu	Phe	Ser	Val	Asn
Leu														
	50					55						60		
Thr	Val	Pro	Glu	Asn	Val	Ser	Gln	Val	Thr	Val	Tyr	Pro	Val	Tyr
Asp														
	65				70				75				80	
Glu	Asp	Tyr	Gly	Leu	Gly	Arg	Leu	Val	Asn	Thr	Ala	Asp	Ala	Ser
Gln														
	85							90				95		
Ser	Ile	Ile	Tyr	Gln	Ile	Val	Asp	Glu	Lys	Gly	Lys	Lys	Met	Leu
Lys														
	100						105					110		
Asp	His	Gly	Ala	Glu	Val	Thr	Pro	Asn	Gln	Gln	Ile	Thr	Phe	Lys
Ala														
	115						120					125		
Leu	Asn	Tyr	Thr	Ser	Gly	Glu	Lys	Lys	Ile	Ser	Pro	Gly	Ile	Tyr
Asn														
	130					135						140		
Asp	Gln	Val	Met	Val	Gly	Tyr	Tyr	Val	Asn					
	145				150									

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 167 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown

- (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: N-terminal

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: E. coli
- (B) STRAIN: CS6

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Ser	Met Leu Lys Lys Ile Ile Ser Ala Ile Ala Leu Ile Ala Gly Thr			
	1	5	10	15
Asn	Gly Val Val Asn Ala Gly Asn Trp Gln Tyr Lys Ser Leu Asp Val			
	20	25	30	
Arg	Val Asn Ile Glu Gln Asn Phe Ile Pro Asp Ile Asp Ser Ala Val			
	35	40	45	
Leu	Ile Ile Pro Val Asn Tyr Asp Ser Asp Pro Lys Leu Asp Ser Gln			
	50	55	60	
Ile	Tyr Thr Val Glu Met Thr Ile Pro Ala Gly Val Ser Ala Val Lys			
	65	70	75	80
Leu	Ala Pro Thr Asp Ser Leu Thr Ser Ser Gly Gln Gln Ile Gly Lys			
	85	90	95	
Lys	Val Asn Val Asn Asn Pro Asp Gln Asn Met Asn Tyr Tyr Ile Arg			
	100	105	110	
Pro	Asp Ser Gly Ala Gly Asn Phe Met Ala Gly Gln Lys Gly Ser Phe			
	115	120	125	

Gly	Val Lys Glu Asn Thr Ser Tyr Thr Phe Ser Ala Ile Tyr Thr Gly		
	130	135	140
Leu	Glu Tyr Pro Asn Ser Gly Tyr Ser Ser Gly Thr Tyr Ala Gly Asn		
	145	150	155
160	Thr Val Ser Phe Tyr Ser Asn		
	165		

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 212 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown
 - (ii) MOLECULE TYPE: peptide
 - iii) HYPOTHETICAL: NO
 - (iv) ANTI-SENSE: NO
 - (v) FRAGMENT TYPE: N-terminal
 - (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: E. coli
 - (B) STRAIN: CS6

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Pro Ser Val Gln Ile Ser Asn Asn Lys Ala Tyr Pro Leu Ile Ile
Gln 20 25 30

	Ser	Asn	Val	Trp	Asp	Glu	Ser	Asn	Asn	Lys	Asn	His	Asp	Phe	Ile
Ala															
	35						40							45	

Lys	50	55	60
-----	----	----	----

Arg	Ile Ile Lys Thr Thr Ile Asn Leu Pro Asp Ser Gln Glu Ser Met			
	65	70.	75	80
Ile	Trp Leu Cys Ile Glu Ser Met Pro Pro Ile Glu Lys Ser Thr Lys			
	85	90	95	
Gly	Asn Arg Lys Glu Gly Arg Thr Asp Ser Ile Asn Ile Ser Ile Arg			
	100	105	110	
Phe	Cys Ile Lys Leu Ile Tyr Arg Pro Ala Ser Val Pro Ser Pro Val			
	115	120	125	
Leu	Asn Asn Ile Val Glu Lys Leu Lys Trp His Lys Asn Gly Lys Tyr			
	130	135	140	
Phe	Val Leu Lys Asn Asn Thr Pro Tyr Tyr Ile Ser Phe Ser Glu Val			
	145	150	155	
160				
Lys	Phe Asp Ser Asp Lys Val Asn Asn Ala Lys Asp Ile Leu Tyr Val			
	165	170	175	
Lys	Pro Tyr Ser Glu Lys Lys Ile Asp Ile Ser Asn Arg Ile Ile Lys			
	180	185	190	
Tyr	Ile Lys Trp Ala Met Ile Asp Asp Ala Gly Ala Lys Thr Lys Leu			
	195	200	205	
	Glu Ser Ile Leu			
	210			

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 802 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: N-terminal

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: E. coli
- (B) STRAIN: CS6

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

	Met Asn Gln Phe Tyr Lys Lys Ser His Tyr Ser Ile Gln Lys His			
Gln				
	1	5	10	15
	Ile Thr Gly Leu Leu Phe Leu Leu Phe Ile Tyr Pro Phe Ser Thr			
Ser				
	20	25	30	
	Tyr Gly Asn Glu Gln Phe Ser Phe Asp Ser Arg Phe Leu Pro Ser			
Gly				
	35	40	45	
	Tyr Asn Tyr Ser Leu Asn Ser Asn Leu Pro Pro Glu Gly Glu Tyr			
Leu				
	50	55	60	
	Val Asp Ile Tyr Ile Asn Lys Ile Lys Lys Glu Ser Ala Ile Ile			
Pro				
	65	70	75	80
	Phe Tyr Ile Lys Gly Asn Lys Leu Val Pro Cys Leu Ser Lys Glu			
Lys				
	85	90	95	
	Ile Ser Ser Leu Gly Ile Asn Ile Asn Asn Asn Asp Asn Thr Glu			
Cys				
	100	105	110	

Val Glu Thr Ser Lys Ala Gly Ile Ser Asn Ile Ser Phe Glu Phe
Ser 115 120 125

Ser Leu Arg Leu Phe Ile Ala Val Pro Lys Asn Leu Leu Ser Glu
Ile 130 135 140

Asp Lys Ile Ser Ser Lys Asp Ile Asp Asn Gly Ile His Ala Leu
Phe 145 150 155

160

Phe Asn Tyr Gln Val Asn Thr Arg Leu Ala Asn Asn Lys Asn Arg
Tyr 165 170 175

Asp Tyr Ile Ser Val Ser Pro Asn Ile Asn Tyr Phe Ser Trp Arg
Leu 180 185 190

Arg Asn Leu Phe Glu Phe Asn Gln Asn Asn Asp Glu Lys Thr Trp
Glu 195 200 205

Arg Asn Tyr Thr Tyr Leu Glu Lys Ser Phe Tyr Asp Lys Lys Leu
Asn 210 215 220

Leu Val Val Gly Glu Ser Tyr Thr Asn Ser Asn Val Tyr Asn Asn
Tyr 225 230 235

240

Ser Phe Thr Gly Ile Ser Val Ser Thr Asp Thr Asp Met Tyr Thr
Pro 245 250 255

Ser Glu Ile Asp Tyr Thr Pro Glu Ile His Gly Val Ala Asp Ser
Asp 260 265 270

Ser Gln Ile Ile Val Arg Gln Gly Asn Thr Ile Ile Ile Asn Glu
Ser

	275	280	285
Thr	Val Pro Ala Gly Pro Phe Ser Phe Pro Ile Thr Asn Leu Met Tyr		
	290	295	300
Lys	Gly Gly Gln Leu Asn Val Glu Ile Thr Asp Ile Tyr Gly Asn Lys		
	305	310	315
320			
Gly	Gln Tyr Thr Val Asn Asn Ser Ser Leu Pro Val Met Arg Lys Ala		
	325	330	335
Ser	Leu Met Val Tyr Asn Phe Ile Ser Gly Lys Leu Thr Lys Lys Asn		
	340	345	350
His	Glu Asp Gly Asp Phe Phe Thr Gln Gly Asp Ile Asn Tyr Gly Thr		
	355	360	365
Phe	Tyr Asn Ser Thr Leu Phe Gly Gly Tyr Gln Phe Ser Lys Asn Tyr		
	370	375	380
Trp	Asn Leu Ser Thr Gly Ile Gly Thr Asp Leu Gly Phe Ser Gly Ala		
	385	390	395
400			
Asn	Leu Leu His Val Ser Arg Ser Asn Phe Lys Asn Lys Asn Gly Tyr		
	405	410	415
Val	Ile Asn Leu Gln Gln Asn Thr Gln Leu Arg Pro Phe Asn Ala Gly		
	420	425	430
Asp	Asn Phe Asp Tyr Ala Tyr Arg Lys Lys Arg Tyr Val Glu Leu Ser		
	435	440	445
	Ile Gly Trp His Gly Asn Leu Tyr Asn Gln Leu Lys Asn Ser Phe		

Ser			
	450	455	460
	Leu Ser Leu Ser Lys Ser Leu Asn Lys Tyr Gly Asn Phe Ser Leu		
Asp	465	470	475
480			
	Tyr Asn Lys Met Lys Tyr Trp Asp Asn Ala Tyr Asp Ser Asn Ser		
Met		485	490
			495
	Ser Ile Arg Tyr Phe Phe Lys Phe Met Arg Ala Met Ile Thr Thr		
Asn		500	505
			510
	Cys Ser Leu Asn Lys Tyr Gln Ser Tyr Glu Lys Lys Asp Lys Arg		
Phe	515	520	525
	Ser Ile Asn Ile Ser Leu Pro Leu Thr Lys Asp Tyr Gly His Ile		
Ser	530	535	540
	Ser Asn Tyr Ser Phe Ser Asn Ala Asn Thr Gly Thr Ala Thr Ser		
Ser	545	550	555
560			
	Val Gly Leu Asn Gly Ser Phe Phe Asn Asp Ala Arg Leu Asn Trp		
Asn		565	570
			575
	Ile Gln Gln Asn Arg Thr Thr Arg Asn Asn Gly Tyr Thr Asp Asn		
Thr		580	585
			590
	Ser Tyr Ile Ala Thr Ser Tyr Ala Ser Pro Tyr Gly Val Phe Thr		
Gly	595	600	605
	Ser Tyr Ser Gly Ser Asn Lys Tyr Ser Ser Gln Phe Tyr Ser Ala		
Ser		610	615
			620

Gly Gly Ile Val Leu His Ser Asp Gly Val Ala Phe Thr Gln Lys
Ala 625 630 635
640

Gly Asp Thr Ser Ala Leu Val Arg Ile Asp Asn Ile Ser Asp Ile
Lys 645 650 655

Ile Gly Asn Thr Pro Gly Val Tyr Thr Gly Tyr Asn Gly Phe Ala
Leu 660 665 670

Ile Pro His Leu Gln Pro Phe Lys Lys Asn Thr Ile Leu Ile Asn
Asp 675 680 685

Lys Gly Ile Pro Asp Gly Ile Thr Leu Ala Asn Ile Lys Lys Gln
Val 690 695 700

Ile Pro Ser Arg Gly Ala Ile Val Lys Val Lys Phe Asp Ala Lys
Lys 705 710 715
720

Gly Asn Asp Ile Leu Phe Lys Leu Thr Thr Lys Asp Gly Lys Thr
Pro 725 730 735

Pro Leu Gly Ala Ile Ala His Glu Lys Asn Gly Lys Gln Ile Asn
Thr 740 745 750

Gly Ile Val Asp Asp Gly Met Leu Tyr Met Ser Gly Leu Ser
Gly 755 760 765

Thr Gly Ile Ile Asn Val Thr Trp Asn Gly Lys Val Cys Ser Phe
Pro 770 775 780

Phe Ser Glu Lys Asp Ile Ser Ser Lys Gln Leu Ser Val Val Asn
Lys 785 790 795
800